PERGAMON

Available online at www.sciencedirect.com

SCIENCE DIRECT®

Vision Research 43 (2003) 2527–2537

Vision Research

www.elsevier.com/locate/visres

Power spectra and distribution of contrasts of natural images from different habitats

Rosario M. Balboa [a], Norberto M. Grzywacz [b,*]

[a] *Departamento de Biotecnología, Facultad de Ciencias, Universidad de Alicante, Apartado de Correos 99, 03080 Alicante, Spain*

[b] *Department of Biomedical Engineering, Neuroscience Graduate Program, and Center for Vision Science and Technology, University of Southern California, University Park OHE 500, Los Angeles, CA 90089-1451, USA*

Received 14 May 2001; received in revised form 4 November 2002

Abstract

Some theories for visual receptive fields postulate that they depend on the image statistics of the natural habitat. Consequently, different habitats may lead to different receptive fields. We thus decided to study how some of the most relevant statistics vary across habitats. In particular, atmospheric and underwater habitats were compared. For these habitats, we looked at two measures of the power spectrum and one of the distributions of contrasts. From power spectra, we analyzed the log–log slope of the fall and the degree of isotropy. From the distribution of contrasts, we analyzed the fall in a semi-log scale. Past studies found that the spatial power spectra of natural atmospheric images fall linearly in logarithmic axes with a slope of about −2 and that their distribution of contrasts shows an approximate linear fall in semi-logarithmic axes. Here, we show that the power spectrum of underwater images have statistically significantly steeper slopes (≈−2.5 in log–log axes) than atmospheric images. The vast majority of power spectra are non-isotropic, but their degree of anisotropy is extremely low, especially in atmospheric images. There are also statistical differences across habitats for the distribution of contrasts, with it falling faster for underwater images than for atmospheric ones. We will argue that these differences are due to the optical properties of water and that the differences have relevance for theories of visual receptive fields. These theories would predict larger receptive fields for aquatic animals compared to land animals.

Keywords: Power spectrum; Contrast; Scale invariance; Natural image; Receptive field; Habitat

1. Introduction

Barlow (1961a, 1961b, 1989) postulated that the function of early receptive fields in the visual system is to eliminate spatial redundancies from the visual code. Several theoreticians (Atick & Redlich, 1990, 1992; McCarthy & Owen, 1996; Srinivasan, Laughlin, & Dubs, 1982 and personal communication) proposed specific receptive-field models that implement Barlow's redundancy ideas. One of these models (Srinivasan et al.'s) was based on the autocorrelation function, which is a statistical technique to quantify image redundancies. The other models used the spatial power spectrum of the image, which is the Fourier transform of the autocorrelation (Wiener, 1964). It was thus important to measure the statistical redundancies, and in particular, the power spectra of natural images. Several investigators made careful measurements of such spectra in images transmitted through air, concluding that these spectra tend to fall with the square of spatial frequency (Burton & Moorhead, 1987; Carlson, 1978; Field, 1987, 1993, 1994; Ruderman & Bialek, 1994; Tolhurst, Tadmor, & Chao, 1992; van der Schaaf & van Hateren, 1996; van Hateren, 1992; McCarthy and Owen, unpublished observation). Images with such power spectra were often called scale-invariant. Based on scale-invariance, the models above predicted specific early receptive-field structures in land animals. Would receptive fields have different structures if natural spectra were different? We investigated power spectra of natural underwater images, hoping that the special optical properties of water

* Corresponding author.

E-mail addresses: rosario.balboa@ua.es (R.M. Balboa), nmg@bmsr.usc.edu (N.M. Grzywacz).

0042-6989/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/S0042-6989(03)00471-1

(high absorption and scattering Mobley, 1995) may make their power spectra different from those of atmospheric images.

In addition, we studied the contrasts distribution in images from both habitats to test whether there are significant differences in these two habitats. The importance of contrasts was that they figured prominently in a more recent theory of early receptive-field function (Balboa & Grzywacz, 2000a; Grzywacz & Balboa, 2002). That theory postulated that the goal of sensory adaptation was to set the internal parameters of the system such as to minimize errors in the performance of specific tasks. An application of the theory to the retina suggested that two of its tasks were the estimation of contrasts and of positions of occluding borders. Because the ease to estimate position depended on the contrast of the border, contrast had a special status in the theory. The distribution of contrasts had been measured in atmospheric images (Balboa & Grzywacz, 2000b; Ruderman & Bialek, 1994; Zhu & Mumford, 1997). It fell with contrasts such that at intermediate frequencies the fall appeared exponential. Another measured contrast-related function was the probability that a border passes through a point given its contrast (Balboa & Grzywacz, 2000b). The sigmoidal shape of this probability function was an essential feature of the application of the theory mentioned above. Balboa and Grzywacz (2000b) showed that this probability function and the distribution of contrasts are related, obviating the need to measure the former.

Some of this work appeared previously in abstract form (Balboa & Grzywacz, 2001).

2. Methods

2.1. Images

The 113 gray-scale images used in this study were taken using three camera types (waterproof and outdoor Kodak MAX cameras loaded with Kodak MAX 800 film, and Quick Snap Outdoor camera loaded with Fuji color Superior X-TRA 800 film). The images were taken from two atmospheric (Point Reyes, California and La Tovara, State of Nayarit, Mexico) and two underwater (John Pennekamp Coral Reef State Park, Florida and Isla de Tabarca, Alicante, Spain) habitats. There were not any human-made objects in the images (Field, 1987, 1993, 1994) and pictures were taken without aiming on any particular subject. After scanning, the images had sizes greater than 500×500 pixels. In Fig. 1, one can observe some typical underwater and atmospheric examples of these images.

Fig. 1. Five natural images used in this work. The figure shows examples of atmospheric images (a3_16 and a2_13) and underwater images (i4_05, i5_05, and i1_10).

2.2. Intensity calibration

To calibrate light intensity, a grid of 16 gray-level rectangles (each rectangle measured 5×3.2 cm^2) with evenly spaced densities was produced. The intensity on the center of each rectangle was measured outdoors in a sunny day, using a spot meter photometer (model UBD 10; Photo Research, Kollmorgen Corporation). We took measurements from 1.5 to 2 m. The measured intensities ranged between 8.7 and 102.7 candelas/m^2 from the darkest to the brightest rectangle. This relatively narrow range of rectangle intensities was sufficient for our purposes, since contrasts in any given image are low (Balboa & Grzywacz, 2000b; Ruderman & Bialek, 1994; Vu, McCarthy, & Owen, 1997; Zhu & Mumford, 1997), including those in this study. We then photographed (from 1.5 m) with each camera the grid of gray-level rectangles. These photographs were scanned at 94.5 dots/mm with an IBM Idea Scan 2000 scanner. An intensity-calibration curve was built for each film by interpolating with a spline the plot of the sixteen photometer-measured intensities versus the corresponding mean scanned gray levels. This curve was applied to each scanned (11.8 dots/mm) natural image to convert its gray levels to light intensity units.

2.3. Power spectrum

The images were rectangular, but the analysis cropped them to the largest, concentric square portion. We used the mixed-radix Fourier-transform algorithm (Nussbaumer, 1982) implemented by MATLAB (MathWorks, Natick, MA) to obtain the spectra from the images.

2.3.1. Calibration of high spatial frequencies

The quality of images was affected by the imperfect optics of the lens of the camera. To measure the quality of the optics, the modulation transfer function (MTF) was used. The MTF of the lens typically expresses losses of high spatial frequencies, which lower the sharpness of edges. The optics of the camera was taken into account by determining the response to thin lines and correcting for the changes produced in the spectrum (Field, 1994). For this calibration, we took pictures of vertical lines with widths of 391 and 730 μm. In two out of six films, the thin line could be resolved and was thus used in the calibration, while in three other films, we could only use the thick line. The sixth film was discarded due to the lack of both lines.

Another problem was that occasionally, the power spectra obtained from our natural images showed vertical and horizontal lines of high spectral energy running from the low to the high spatial frequencies. Vertical lines occurred at low horizontal frequencies and horizontal lines occurred at low vertical frequencies. As footnote 1 explains, these lines were artifacts of the square truncation of images. [1] The standard way to deal with such artifacts is to window the images with a circularly symmetric smooth function that tapers to zero at the boundary. Unfortunately, this windowing method would not work for us. From the convolution theorem, the Fourier transform of such a windowed image is the convolution of the Fourier transform of the image with Fourier transform of the window. Consequently, windowing the images like this would change the slopes of their power spectra, invalidating our analysis. We thus developed an alternate method to deal with the power-spectrum truncation lines.

We used a statistical procedure taking into account that these lines extended to spatial frequencies so high that only noise was present at the rest of the spectrum. Hence, the procedure discarded lines at low spatial frequencies that were outliers relative to high spatial-frequency noise. To estimate this noise, we selected two squares of 50×50 pixels, one to eliminate lines that were horizontal outliers and one to eliminate lines that were vertical outliers. If X and Y were the horizontal and vertical sizes of the spectrum, the upper-left corners of the squares were placed at positions $(1, Y/2 + 10)$ and $(X/2 + 10, Y + 1 - 50)$ respectively. This placement assumed that (1,1) was the upper-left corner of the image. To eliminate a horizontal line of the power spectrum between $Y/2 - 9$ and $Y/2 + 9$, we computed the mean of the logarithm of power in that line between positions 1 and 50. Only if this mean was within the central 99% of a Gaussian with the same mean log-power and variance as in the square, the line remained. For symmetry, if the analysis eliminated a horizontal line, it also made sure to eliminate all the horizontal and vertical lines equidistant to the center of the spectrum. A similar procedure (using the second square) eliminated outlier vertical lines. Therefore, the number of eliminated lines was always the same, making the final spectrum a square. No more than seventeen rows or columns were ever eliminated using this statistical procedure. This procedure did not affect the estimation of slopes of power spectra, because the remaining lines maintained their true energies and spatial frequencies.

[1] These lines were artifacts, since they disappeared when circular patches of the images (with radii equal to 80% of the image sides) were cropped and embedded concentrically with square windows to obtain the spectra. (The spectra thus obtained were similar to the spectra obtained without cropping, except for the lines. Thus, our results would be the same if we cropped the images or used the line-removal technique described in the text.) And in contrast, artifacts occurred for artificial circularly symmetric Gaussian images truncated with square windows of sides smaller or equal than the Gaussian's standard deviation (SD). Mathematical analysis suggested that all the artifacts discussed in this footnote emerged from truncation of images with long-range positive autocorrelation.

In other words, our procedure simply removes those spectral lines that have the artifact in a statistically significant way. To perform the relevant statistics, we measure the noise in regions of high spectral frequencies. Because lines are removed statistically, those lines that do remain (which are the vast majority of them) maintain their spectral powers untouched. The remaining lines thus allow an unbiased calculation of the slopes.

2.3.2. Calculation of slope and isotropy

We subtracted the logarithm of the power spectrum of the vertically averaged line from the logarithm of the rotationally averaged power spectrum of the image. The slopes of the fall of the resulting power spectra were estimated from a linear regression of the logarithm of the power as a function of the logarithm of spatial frequency. To alleviate the effect of high-frequency noise on the slope, we constrained the regression to frequencies lower than 100 cycles/image.

We were also interested to test whether the power spectrum was isotropic and if so, whether this property held across habitats. Our goal was to detect possible anisotropies in individual images rather than to test whether particular kinds of absolute angles of anisotropies (for instance, vertical and horizontal) dominated across images. Measuring such angles would have required using a constant positioning of the camera, which would be hard for underwater images. To quantify anisotropy in each image, we applied the following procedure: Before calculating the rotational mean of each spectrum (necessary to estimate slopes), the largest possible concentric circular patch was cropped and its logarithm integrated in 72 non-overlapping 5° sectors ($72 \times 5° = 360°$), with the first sector being horizontal. We then arranged the concentric patches in 24 samples of three contiguous sectors, with the first horizontal sector belonging to the same sample as its neighbors. Finally, we tested with one-way analysis of variance whether the 15° samples came from populations with identical mean integrated log-powers. Differences across these populations indicated that the spectrum was statistically significantly anisotropic. The degree of anisotropy in each image was evaluated with the coefficient of variation (ratio SD over mean) across the 72 sectors (Burgi & Grzywacz, 1998).

2.4. Distribution of contrasts

The definition of local contrast for an image point at Position (i, j) is the ratio between the absolute value of the two-dimensional gradient of intensity ($|\nabla I|_{i,j}$) and the local mean intensity ($\overline{I}_{i,j}$—Balboa & Grzywacz, 2000b). The discrete approximation for the absolute value of the gradient of intensity is

$$|\nabla I|_{i,j} = \sqrt{\left(\frac{I_{i+1,j} - I_{i-1,j}}{2\Delta x}\right)^2 + \left(\frac{I_{i,j+1} - I_{i,j-1}}{2\Delta y}\right)^2},$$

where $\Delta x = \Delta y$ are the distances between neighbor pixels. Without loss of generality we set $\Delta x = \Delta y = 1$. In turn, the mean local intensity is

$$\overline{I}_{i,j} = \frac{I_{i,j+1} + I_{i,j-1} + I_{i+1,j} + I_{i-1,j}}{4}.$$

For each image, local contrasts for all points were arranged in frequency histograms and the logarithm of the absolute frequency as a function of contrast was linearly fitted with the resistant-line-from-three-groups algorithm (Mosteller & Tukey, 1983). This type of fit came from Robust statistics, and thus was insensitive to outlier deviations from the fit. From the linear fits, slopes were extracted. The goodness of the fit for each distribution of contrasts was measured with the median absolute deviation (MAD) ratio (Sprent, 1993). Contrasts yielding MAD ratios larger than 5 were considered outliers, that is, deviating significantly from the prediction of the resistant line.

3. Results

3.1. Power spectrum

Power spectra are important image statistics for theories that propose that the function of the early visual system is to eliminate spatial redundancies in the image (see Section 1). In this section, we compare power spectra from atmospheric and underwater images. After elimination of truncation and high-frequency errors, power spectra of natural calibrated images look like those in Fig. 2. Not surprisingly, power spectra fall as the spatial frequency increases. The evidence for this is that the centers of the spectra are brighter than the periphery. One can also observe that power spectra of natural images seem mostly isotropic (specially images i1_10 and a2_13). However, one can see a degree of anisotropy in almost all images (for example, images i4_05 and a3_16 have diagonal axes of anisotropy between 0° and 90°).

From spectra as shown in Fig. 2, we plotted their energy averaged over orientations as a function of spatial frequency (Fig. 3). In a log–log scale, these plots were approximately straight lines. Although the slopes of these lines were often close to −2 (for instance, images i1_10 and a3_16), these slopes could also differ from this value (images i4_05 and a2_13). For instance, one could occasionally also see a small positive-second-derivative bending of the power spectrum at high spatial frequencies (images i1_10 and a2_13). This was probably due to image noise, which is only significant at high frequencies, where the signal becomes small.

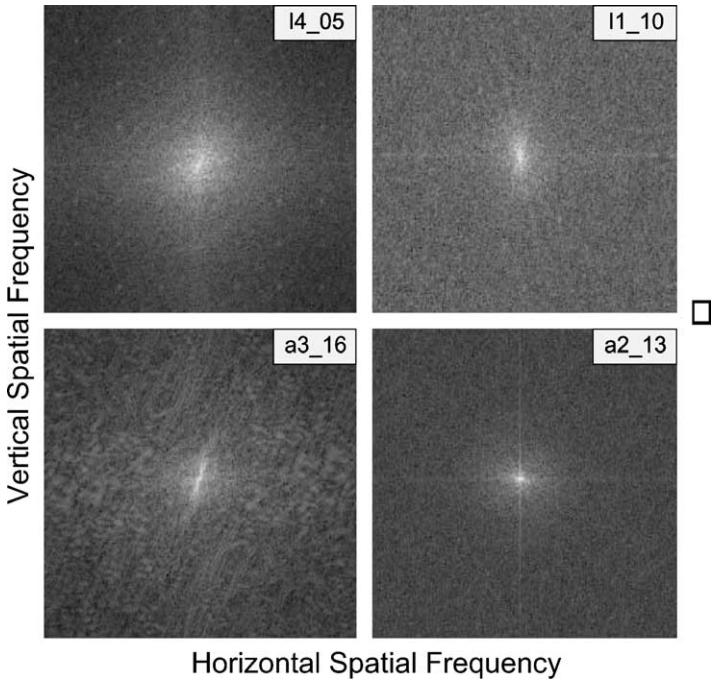

Fig. 2. Power spectra of four images shown in Fig. 1. The horizontal and vertical orientations correspond to the horizontal and vertical components of the spatial-frequency vector respectively, and the logarithm of power is represented by the whiteness of the pixels. Absolute spatial frequency is proportional to the distance from the center of the spectrum. (We omitted scales of power and spatial frequency for the sake of clarity, as the intention of the figure is to show power spectra qualitatively.) Each spectrum has a different level of low-frequency truncation and high-frequency-error elimination (see Section 2). The power always falls with spatial frequency. In addition, the spectra may be more (i4_05 and a3_16) or less (i1_10 and a2_13) anisotropic.

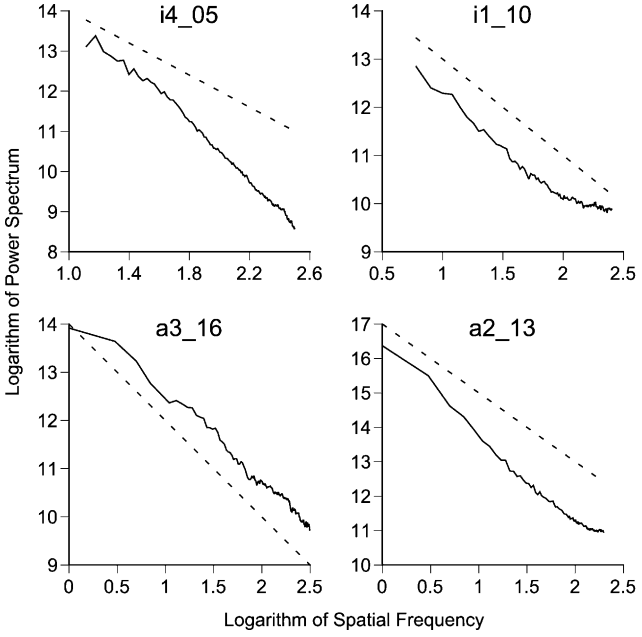

Fig. 3. The power spectra of Fig. 2 averaged over orientations. The dashed line represents the log–log slope of −2. In all cases, the power spectra fall approximately linearly with the spatial frequency in log–log axes (solid line). The slope of the fall hovers around −2, but it is not always so.

From plots as shown in Fig. 3, we measured the log–log slopes of power spectra of underwater and atmospheric images (Fig. 4). [2] The slopes from underwater images (-2.5 ± 0.06; standard error, SE) were statistically significantly steeper than those from atmospheric images (-2.0 ± 0.04 – one-sided Mann-Whitney test, $U = 615$, $z = -5.41$, $p < 10^{-4}$). Underwater, but not

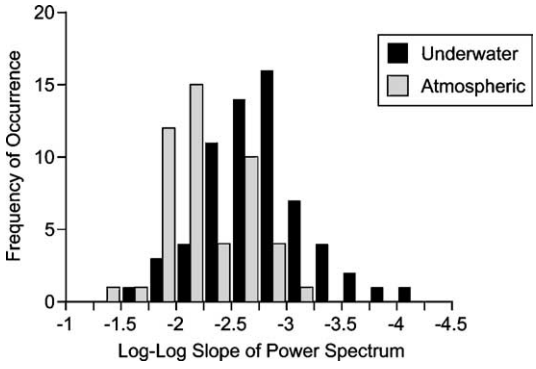

Fig. 4. Histograms of the log–log slopes of the power spectra of underwater and atmospheric images. Underwater images produce slopes more negative than atmospheric ones. The figure also emphasizes the wide variability in slopes, as reported previously (Tadmor & Tolhurst, 1994; van der Schaaf & van Hateren, 1996).

[2] To alleviate the effect of high-frequency noise on the calculation of slopes, we constrained this calculation to frequencies up to 100 cycles/image (see Section 2).

atmospheric images had log–log slopes significantly different from −2 (two-sided Wilcoxon Rank test, $S = 129$, $n = 64$, $Z = -6.09$, $p < 10^{-4}$).

Hence, although our atmospheric power-spectrum slopes conform with the atmospheric slopes reported elsewhere (see Section 1), our underwater slopes do not.

Besides slopes, the analysis also looked at anisotropies on the power spectrum. Fig. 5 illustrates these anisotropies in two images by plotting the integrated power in 15° sectors at various orientations. The F statistic of the analysis of variance applied to data like those in Fig. 5 showed that there were statistically significantly anisotropies in the majority of the power spectra (Fig. 5, image a3_16). However, the anisotropies were weak and sometimes absent (image i5_05). The vast majority of images (107 images out of 113) showed anisotropy at the $p < 0.01$ significance level.

One way to realize that the anisotropies were small was to pay attention to the powers in the vertical axes of Fig. 5. It was not difficult to see that variations of power were a small fraction of the mean power. In Fig. 6, we quantified this smallness by using the coefficient of variation. This coefficient revealed that the spectral energy in any given orientation did no deviate by much more than 2% from the mean energy across orientations. Nevertheless, these deviations were sufficient to tell that the degree of anisotropy was larger for underwater (0.014 ± 0.0011; SE) images than for atmospheric ones (0.008 ± 0.001 – two-sided-Mann-Whitney test, $U = 830.5$, z corrected for ties $= -4.24$, $p < 10^{-4}$). When inspected images with relatively high coefficients of variation, the preferred orientation in plots like those in Fig. 5 could occur in any angle. And in the 13 images (12 underwater and one atmospheric) with a coefficient of variation coefficient larger than 0.02, we found more than one preferred orientation (usually two or four).

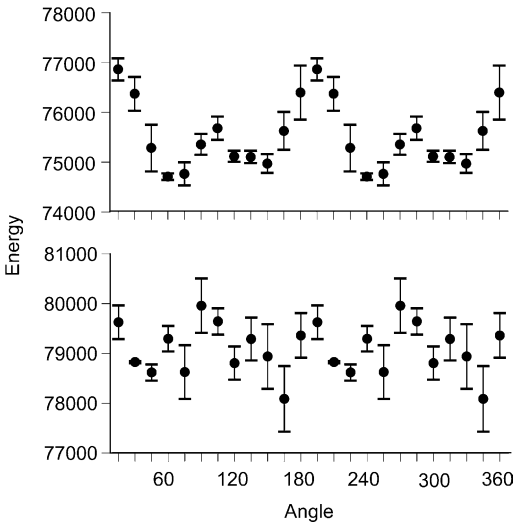

Fig. 5. Integrated power as a function of orientation for spectra as shown in Fig. 2. Each graph shows the power integrated in non-overlapping 5° sectors and averaged over every three consecutive ones for an overall 15° sector. The data points are these averages and the error bars are the standard deviations. There are anisotropies in the spectra (for instance, in a3_16—top graph), but these anisotropies are small (note the absolute values of the vertical axes) and sometimes not very apparent (i5_05—bottom graph).

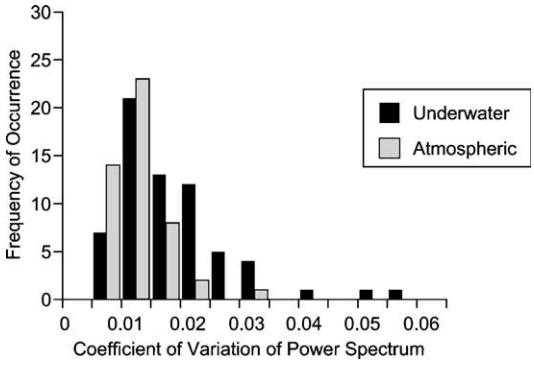

Fig. 6. The coefficient of variation of the means of the 15° sectors in both habitats. This statistic was extracted from the analysis of variance applied to data like those in Fig. 5. The degree of anisotropy was statistically higher for underwater images than for atmospheric images.

3.2. Distribution of contrasts

Besides power spectra, another important image statistics for theories of the early visual system is the distribution of contrasts. This distribution is essential for theories that emphasize border localization and estimation of contrast (see Section 1).

Fig. 7 shows distributions of contrasts for six natural images. The three upper panels come from underwater images, while the lower three panels come from atmospheric ones. The general trend of this distribution for all images is that of an exponential fall at intermediate contrasts (Fig. 7, images i1_14 and i1_10). However, atmospheric distributions tend to fall in two faster-than-exponential phases at low and high contrasts. Moreover, at these contrasts, the majority of the images, atmospheric or underwater, present statistically significant deviations from the intermediate-contrast linear trend, which are quantified as outlier residuals between the data curve and the fitted line (asterisks shown in Fig. 7). In 8 out of 113 images (five atmospheric and three underwater), we observe prominent, narrow peaks (images a3_16 and a2_13 in Fig. 7), which represent strong deviations from the general linear trend. These peaks appear either at very low (Fig. 7, image a3_16), or at intermediate and very high contrasts (Fig. 7, image a2_13).

Fig. 8 summarizes the slopes of the semi-log, robust, linear fits distributions of contrasts of underwater and atmospheric habitats. There is a statistically significantly

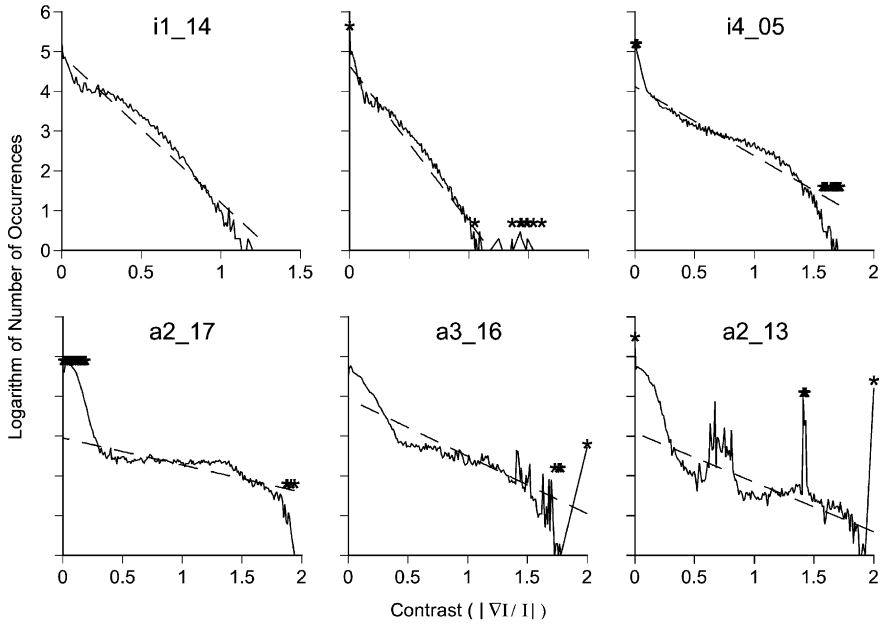

Fig. 7. Distribution of contrasts from six images in a semi-log scale. The solid lines are the raw distribution of contrasts and the dashed lines are the linear fits using the resistant-line-from-three-groups method (see Section 2). (This type of fit comes from Robust statistics, and thus is insensitive to outlier deviations from the fit.) Underwater images (i1_14, i1_10, and i4_05) have steeper slopes than atmospheric ones (a2_17, a3_16, and a2_13). At very low and very high contrasts, the linear trend (representing an exponential fall in the semi-log scale) is violated, with statistically significant violations labeled with asterisks in this figure. In a very few cases, there are also violations of the linear trend with narrow peaks at intermediate contrasts.

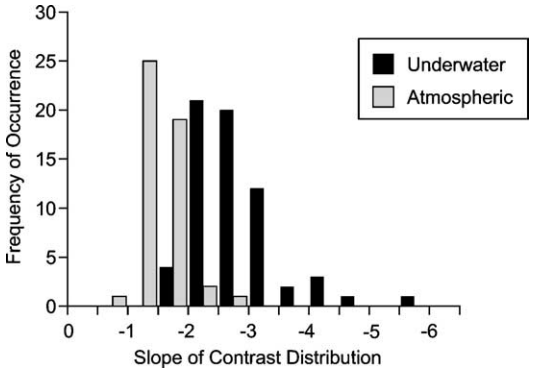

Fig. 8. Histograms of the slopes of the fall of the distribution of contrasts in a semi-log scale for all images. Underwater images produce slopes statistically significantly steeper than atmospheric images.

difference between underwater and atmospheric slopes of the distribution of contrasts. Underwater images have steeper slopes that atmospheric ones (one-sided-Mann-Whitney test, $U = 63$, $z = -8.66$, $p < 10^{-4}$). Finally, to analyze in which habitat the exponential trend is most often violated, we analyze the outliers coming from the residuals between the data curve and the resistant line. Atmospheric images have statistically significantly more outliers than underwater images when we group outliers at low (2×2 contingency table, χ^2 with continuity correction $= 30.3$, $p < 10^{-4}$) and at the rest of the contrast range ($\chi^2 = 16.5$, $p < 10^{-4}$).

4. Discussion

The log–log slopes of the power spectra were statistically significantly more negative than −2 for underwater images, but hovered around −2 for atmospheric images (Figs. 3 and 4). There were statistically significant anisotropies in the vast majority of the images (Fig. 5), but their degrees of anisotropy (quantified by the coefficient of variation) were small (Fig. 6). As for the distribution of contrasts, we observed that at intermediate contrasts there was a general linear trend in a semi-log scale in all images (Fig. 7). At low and high contrasts, the fall was faster than linear, especially for atmospheric images. Moreover, the intermediate-contrast slopes of underwater images were statistically steeper than the slopes of atmospheric images (Fig. 8). Finally, the deviations from this linear trend were more prominent in atmospheric images than in underwater ones.

4.1. Limitations

Before discussing the implications of these results, we address some of the limitations of our conclusions on power spectra and contrasts. In this work, as well as in the rest of the literature, high spatial-frequency problems with power spectra have been partially corrected by the line-spread-function method (Field, 1987, 1993, 1994). To understand why this correction is only partial, consider the physics of the MTF of a lens (equivalent to

the Fourier transform of the line spread function). The MTF is a function of the distance to the optical axis of the lens (Landt, 1998). Moreover, the MTF is a function of the angle between the grid used to measure the MTF and the line connecting the center of the lens to the point of entry of the grid rays (Landt, 1998). Hence, strictly speaking, it is not sufficient to correct power spectra of images by presenting just one line perpendicular through the optical axis of the objective of the camera. That others and we did that is only a first-order correction to the spectrum. This paper improves the line-spread-function method by clipping some high frequencies in the computation of the slope of the power spectrum (see footnote 2). A price one pays for this clipping is that the range of usable spatial frequencies is narrower, making the slopes more sensitive to noise.

There may be also a limitation stemming from our definition of contrast. This variable is defined here locally, namely, its value on a point depends only on the intensities at this point and its immediate neighborhood. In some other papers, contrast is more global as it is the discrepancy from the mean intensity of the image (Vu et al., 1997). At first, this difference does not appear to generate problems as our results follow roughly what was observed previously by other authors (Carlson, 1978; Tolhurst et al., 1992; van der Schaaf & van Hateren, 1996). However, we found in a small set of images unexpected narrow peaks at intermediate and high contrasts, which were not reported previously. These peaks can be explained because of the local normalization of the gradient in our definition of contrast (see Section 2.4). To understand this conclusion, consider a bright object occluding a darker one (or vice-versa). It is possible in natural images for such objects to have a large intensity difference. To measure the local contrast at a point near the occlusion border between these objects, we use the four points that are neighbor to this point (see equations in Section 2.4). Because of the large intensity difference, one would find that one, two, or three of these neighbor points are bright and have similar intensities. In contrast, the remaining neighbor points are much darker. Consequently, these low-intensity points make negligible contributions to the gradient and mean intensities used in the definition of local contrast. Moreover, the contrast is the ratio between the gradient and the mean intensity. Hence, the absolute value of intensity of the bright points becomes irrelevant if it is similar across them. This is because the ratio factors out this value of intensity. Thus, occlusion borders often yields one of three values of ratios, that is, stereotypical contrasts, such as observed in panel a2_13 of Fig. 7. This same argument without the normalization or with global contrast would smear the contrast values and the peaks in the distribution would be less evident. We use a local definition of contrast, since local contrast is the most relevant variable for edge localization, a major ingredient of a recent theory of early visual receptive fields (Balboa & Grzywacz, 2000a; Grzywacz & Balboa, 2002).

4.2. Power spectrum

A result shared by this paper and several others in the literature is that the log–log slope of the fall with frequency of the spectra of natural atmospheric images is approximately −2 (Figs. 2 and 3, but see Tolhurst et al., 1992). Slopes around −2 have typically been attributed to self-similarity in images (Field, 1987, 1993, 1994; Mandelbrot, 1977; Ruderman, 1997). The intuition for self-similarity is that the distribution of sizes of features and intensities in images is such that if one were to get closer to the image (thus increasing the sizes and intensities), then the distributions would remain the same. However, a recent paper refuted some of the arguments for self-similarity and pushed an alternate hypothesis (Balboa, Tyler, & Grzywacz, 2001—for further discussion on this matter see Grzywacz, Balboa, & Tyler (2002) and Ruderman (2002)). This hypothesis is based on natural images having luminance edges, whose spectra fall as frequency squared (Carlson, 1978; Tolhurst et al., 1992). Balboa et al. showed that this hypothesis is consistent with frequency-squared spectra under a wide variety of distributions of sizes. They provided two reasons: First, for every frequency, the log–log slope of the rotationally averaged power spectrum of an image is the weighted mean of the log–log slopes from the independent regions of the image formed by objects occluding one another. Second, the log–log slopes of the spectrum envelope for a constant-intensity two-dimensional region are between 0 and −3 for frequencies corresponding to periods much larger and much smaller than the size of the region respectively. Therefore, it is not surprising that natural images have log–log slopes between −1.5 and −3.

Why are the slopes steeper than −2 for underwater images? This is a puzzling problem, since explanations given for −2 slopes in atmospheric environments seem to apply to underwater ones. There is no reason to believe that the distributions of underwater object sizes are different from those in land, supporting models of self-similarity. Furthermore, underwater objects occlude each other like land objects, thus being consistent with the Balboa et al.'s hypothesis. Therefore, because underwater objects obey the conditions for −2 slopes, the violation of these slopes in underwater images must result from the medium connecting objects to images. In other words, underwater images have slopes steeper than −2 because of the optics of water. Three aspects of this optics can cause steeper slopes: First, water has a high degree of scattering for visible light (Mobley, 1995), which would be equivalent to a low-pass filtering. Second, water has a high degree of absorption, especially in

the middle and low wavelengths (Mobley, 1995). Hence, underwater images have a relative lack of distant and thus, small objects, missing power in high spatial frequencies. Third, wind-induced turbulence near the surface of the water causes spatial and temporal anomalies in the refracting index of water. Consequently, any planar wave front of light passing through this turbulence will experience phase distortions before it reaches the camera. The phase distortions blur the images, resulting in a loss of high spatial frequencies. These three forms of optical blurring are especially bad in muddy environments, or in waters with a high concentration of particles and microscopic life, such as the Sargasso Sea (Mobley, 1995).

4.3. Isotropy

Another property of power spectra of natural images is that their degree of anisotropy is significant but small ($\approx$1%). This is equivalent to say that on average there is little difference between the amount of "stuff" at various orientations. One implication is that models of receptive fields are safe in assuming isotropy, at least as a first approximation (Atick & Redlich, 1992; Balboa & Grzywacz, 2000a; Grzywacz & Balboa, 2002; Srinivasan et al., 1982). The anisotropy, however, is statistically significant as noted by others (van der Schaaf & van Hateren, 1996). Although some of the anisotropy is due to the vertical and horizontal biases induced by gravity (van der Schaaf & van Hateren, 1996; Zetsche, Barth, & Wegman, 1993), axes of anisotropy did not always follow these biases.

We also observed that underwater images are more anisotropic than atmospheric ones. This result seems to go against the intuition that atmospheric images have more biases due to gravity given the oblique angle of most underwater images. To try to understand this surprising result, we looked at the images with the highest coefficients of variation, namely, with the highest degree of anisotropy. We found that underwater images with high coefficients of variation have just a few small oriented details (for instance, sea waves seen from below or fish stripes) against a smooth background (due to light absorption). In contrast, atmospheric images have much more textured backgrounds, masking the anisotropy coming from trees or grass. Therefore, that the anisotropy is higher in underwater images than in atmospheric ones has more to due with the general "smoothness" of the underwater background.

4.4. Distribution of contrasts

As it was the case for the power spectrum, a model for the distribution of contrasts in atmospheric images exists (Balboa & Grzywacz, 2000b). This model postulates that at low contrasts, this distribution is dominated by quantal noise in the detection of the image (Balboa & Grzywacz, 2000a; Grzywacz & Balboa, 2002). In turn, at high contrasts, occluding borders would control the behavior. Analysis showed that if the partial distributions of contrast due to noise and occlusions covered separate ranges of contrast, then the resulting distribution could be nearly exponential at intermediate contrasts. However, at low and high contrasts, one would instead observe faster-than-exponential falls.

We find differences between the distribution of contrasts in underwater and atmospheric habitats. These results are consistent with the model that we just described for contrast in atmospheric images. The differences between underwater and atmospheric habitats can be explained using this model by making the quantal-noise and occluding-border distributions further apart in the case of atmospheric images, and closer for underwater images. Again, optics is probably the reason for these differences as it causes blurred edges and more noise in underwater images due to light absorption, water movement, and light scattering. Noise would shift the quantal-noise distribution to the right, whereas blurring would shift the occluding-border distribution to the left, making the overall contrast distribution smoother.

4.5. Implications for the visual system

The retinal circuitry in animals from underwater habitats might be different from those from atmospheric habitats to compensate for discrepancies in contrasts and at high-spatial frequencies (Mobley, 1995). For related reasons, other habitats might lead to different retinal circuitry. Interesting examples include foggy habitats in which pigeons and seabirds must fly for hours or days (Sinclair, 1985), turbid lakes and rivers, and Arctic and Antarctic environments. And there are even different underwater habitats depending on factors such as the organic particles in the water (Mobley, 1995).

Theories in the literature would predict larger lateral-inhibition extents in animals from underwater animals than in those from atmospheric habitats. In previous work, we analyzed two such theories, which use power spectrum and contrast as the relevant statistics from natural images (Balboa & Grzywacz, 2000b). In Fig. 9, we summarize the predictions of these theories. Fig. 9A shows the predictions of the predictive coding theory (Srinivasan et al., 1982), which proposes that retinal lateral inhibition encodes light intensities such that their wide range fits in the narrow dynamic range of neurons. In turn, Fig. 9B shows the predictions of the maximal decorrelation theory (Atick & Redlich, 1992). This theory proposes that the goal of retinal lateral inhibition is to send to the rest of the brain a signal that is maximally decorrelated in space. Both theories behave similarly

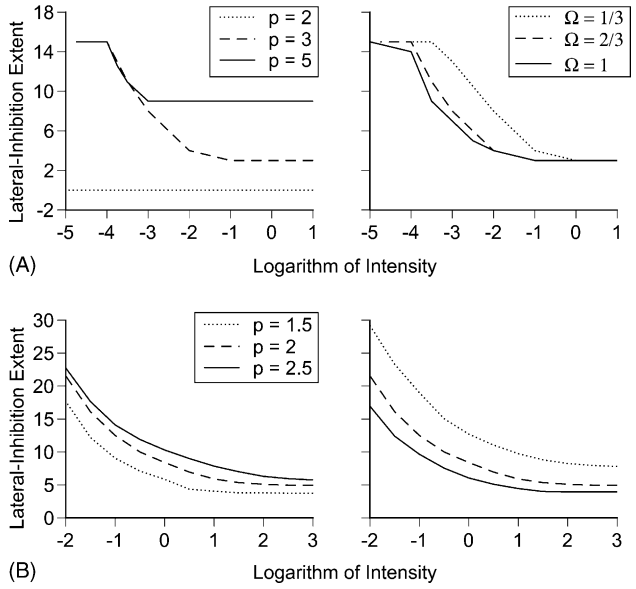

Fig. 9. A: Results of the predictive coding theory, parametric on p (log–log slope of the power spectrum) and Ω (mean contrast). The dashed line is the standard curve ($\Omega = 2/3$, $p = 3$). The other curves in each graph correspond to modulations from the standard curve of only the indicated parameter. For both parameters, the lateral-inhibition extent falls with the mean intensity. This fall is larger when p is smaller and the extent becomes infinitesimally small for $p = 2$ (left panel). As for the mean contrast, the fall starts at lower intensities when Ω is larger (right panel). B: Results of the maximal decorrelation theory. The conventions are as in A and the results are similar. The lateral-inhibition extent falls as the mean intensity increases, as p decreases, and as Ω increases.

parametric on contrasts and spatial-frequency spectra of the images. Increasing the log–log slope of the power spectrum (p) produces larger receptive fields (Fig. 9, left panels). A physical explanation for this effect is that with the loss of high-spatial frequencies, the retinal filters must be restricted to lower frequencies (or wider receptive fields). On the other hand, increasing the mean contrast (Ω) reduces the lateral-inhibition extent (Fig. 9, right panels). The reason for this reduction is that high contrasts mean good signals. Consequently, there is no need to enlarge the receptive field to improve the signal-to-noise ratio, on the contrary. It follows that the two aforementioned theories predict larger receptive fields for underwater images, since these images have larger power-spectra slopes and lower contrasts. A third theory, the local minimal asperity theory (Balboa & Grzywacz, 2000a; Grzywacz & Balboa, 2002), also makes the same prediction, although not from the power spectrum. In this case, the enlargement of receptive fields in aquatic habitats is due to the lack of small objects in underwater images because of turbulence, scattering, and light absorption.

In conclusion, our findings that statistics of natural images vary across habitats may have strong implications for the visual system. Theories about the role of early receptive fields that use these statistics predict strong habitat-dependent differences. Hence, one should be careful when generalizing across habitats the conclusions from studies of natural statistics from restricted habitats.

Acknowledgements

We thank Joaquín de Juan for support while R.M.B. was working in the summer of 2000 at the Department of Biotechnology at the University of Alicante, Spain. We also thank Sonia Mendoza from Matro, Guadalajara, México, for suggesting the inclusion of the La Tovara habitat. Thanks are also due to David K. Fisher for invaluable computer assistance. Finally, we thank the Smith–Kettlewell Institute, where we performed part of this work. This work was supported by National Eye Institute Grants EY08921 and EY11170 to N.M.G., and by a grant from the Programa Ramón y Cajal of the Ministerio de Ciencia y Tecnología of Spain to R.M.B.

References

Atick, J. J., & Redlich, A. N. (1990). Towards a theory of early visual processing. *Neural Computation, 2*, 308–320.

Atick, J. J., & Redlich, A. N. (1992). What does the retina know about natural scenes? *Neural Computation, 4*, 196–210.

Balboa, R. M., & Grzywacz, N. M. (2000a). The minimal-local asperity hypothesis of early retinal lateral inhibition. *Neural Computation, 12*, 1485–1517.

Balboa, R. M., & Grzywacz, N. M. (2000b). Occlusions and their relationship with the distribution of contrasts in natural images. *Vision Research, 40*, 2661–2669.

Balboa, R. M., & Grzywacz, N. M. (2001). Power spectra and distribution of contrasts of natural images from different habitats. *Investigative Ophthalmology and Visual Science, 42*, S615.

Balboa, R. M., Tyler, C. W., & Grzywacz, N. M. (2001). Occlusions contribute to scaling in natural images. *Vision Research, 41*, 964–995.

Barlow, H. B. (1961a). Possible principles underlying the transformation of sensory messages. In W. A. Rosenblith (Ed.), *Sensory communication*. Cambridge, MA: MIT Press.

Barlow, H. B. (1961b). The coding of sensory messages. In W. H. Thorpe, & O. L. Zangwill (Eds.), *Current problems in animal behaviour*. Cambridge University Press.

Barlow, H. B. (1989). Unsupervised learning. *Neural Computation, 1*, 295–311.

Burgi, P.-Y., & Grzywacz, N. M. (1998). A biophysical model for the developmental time course of retinal orientation selectivity. *Vision Research, 38*, 2787–2800.

Burton, C. J., & Moorhead, I. R. (1987). Color and spectral structure in natural scenes. *Applied Optics, 26*, 157–170.

Carlson, C. R. (1978). Thresholds for perceived image sharpness. *Photographic Science and Engineering, 22*, 69–71.

Field, D. J. (1987). Relations between the statistics of natural images and the response properties of cortical cells. *Journal of the Optical Society of America, 4*, 2379–2394.

Field, D. J. (1993). Scale-invariance and self-similar 'wavelet' transforms: An analysis of natural scenes and mammalian visual systems. In M. Farge, J. Hunt, & Vassilicos (Eds.), *Wavelets,*

fractals and Fourier transforms: New developments and new applications (pp. 151–191). Oxford: University Press.

Field, D. J. (1994). What is the goal of sensory coding? *Neural Computation, 6*, 559–601.

Grzywacz, N. M., & Balboa, R. M. (2002). A Bayesian framework for sensory adaptation. *Neural Computation, 14*, 543–559.

Grzywacz, N. M., Balboa, R. M., & Tyler, C. W. (2002). A reply to a letter to the editor by Ruderman. *Vision Research, 42*, 2803–2805.

Landt, A. (1998). Lenses for 35 mm photography. *The KODAK Workshop Series*. Rochester, NY: Silver Pixel Press.

Mandelbrot, B. B. (1977). *Fractals: Form, chance, and dimension*. San Francisco: W.H. Freeman.

McCarthy, S. T., & Owen, W. G. (1996). Preferential representation of natural scenes in the salamander retina. *Investigative Ophthalmology and Visual Science, 37*, S674.

Mobley, C. (1995). Handbook of optics. In M. Bass (Ed.), *Fundamentals, techniques & design* (Vol. 1). New York, NY: McGraw Hill.

Mosteller, F., & Tukey, J. W. (1983). Understanding robust and exploratory data analysis. In D. C. Hoaglin (Ed.), *Wiley sSeries in probability and mathematical statistics*. New York: John Wiley & Sons Inc.

Nussbaumer, H. J. (1982). *Fast Fourier transform and convolution algorithms*. New York, NY: Springer-Verlag.

Ruderman, D. L. (1997). Origins of scaling in natural images. *Vision Research, 37*, 3385–3389.

Ruderman, D. L. (2002). Letter to the editor. *Vision Research, 42*, 2799–2801.

Ruderman, D. L., & Bialek, W. (1994). Statistics of natural images: scaling in the woods. *Physics Review Letter, 73*, 814–817.

Sinclair, S. (1985). *How animals see: other visions of our world*. London: Croom Helm.

Sprent, P. (1993). *Applied nonparametric statistical methods*. London: Chapman & Hall.

Srinivasan, M. V., Laughlin, S. B., & Dubs, A. (1982). Predictive coding: a fresh view of inhibition in the retina. *Proceedings of the Royal Society of London B, 216*, 427–459.

Tadmor, Y., & Tolhurst, D. J. (1994). Discrimination of changes in the second-order statistics of natural and synthetic images. *Vision Research, 34*, 541–554.

Tolhurst, D. J., Tadmor, Y., & Chao, T. (1992). Amplitude spectra of natural images. *Ophthalmic and Physiological Optics, 12*, 229–232.

van der Schaaf, A., & van Hateren, J. H. (1996). Modeling the power spectra of natural images: Statistics and information. *Vision Research, 36*, 2759–2770.

van Hateren, J. H. (1992). Theoretical predictions of spatiotemporal receptive fields of fly LMCs, and experimental validation. *Journal of Comparative Physiology A, 171*, 57–170.

Vu, T. Q., McCarthy, S. T., & Owen, W. G. (1997). Linear transduction of natural stimuli by dark-adapted and light-adapted rods of the salamander, *Ambystoma tigrinum*. *Journal of Physiology, 505.1*, 193–204.

Wiener, N. (1964). *Time Series*. Cambridge, MA: MIT Press.

Zetsche, C., Barth, E., & Wegman, B. (1993). The importance of intrinsically two-dimensional image features in biological vision and picture coding. In A. B. Watson (Ed.), *Digital images and human vision*. Cambridge: MIT Press.

Zhu, S. C., & Mumford, D. (1997). Prior learning and Gibbs reaction-diffusion. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 19*, 1236–1250.